

SEQUENCE LISTING

<110> Sinskey, Anthony J.
Lessard, Philip A.
Willis, Laura B.

<120> Pyruvate Carboxylase from *Corynebacterium glutamicum*

<130> 1533.0790002

<150> US 09/677,575

<151> 2000-10-03

<150> US 09/220,081

<151> 1998-12-23

<160> 39

<170> PatentIn Ver. 2.0

<210> 1

<211> 3621

<212> DNA

<213> *Corynebacterium glutamicum*

<220>

<221> CDS

<222> (199)..(3621)

<400> 1

```

tggggcgggg ttagatcctg ggggggtttat ttcattcact ttggcttgaa gtcgtgcagg 60
tcaggggagt gttgcccga aacattgaga ggaaaacaaa aaccgatgtt tgattggggg 120
aatcgggggt tacgatacta ggacgcagtg actgctatca cccttggcgg tctcttggtg 180
aaaggaataa ttactcta gtg tcg act cac aca tct tca acg ctt cca gca 231
                Met Ser Thr His Thr Ser Ser Thr Leu Pro Ala
                  1                      5                      10
ttc aaa aag atc ttg gta gca aac cgc ggc gaa atc gcg gtc cgt gct 279
Phe Lys Lys Ile Leu Val Ala Asn Arg Gly Glu Ile Ala Val Arg Ala
                  15                      20                      25
ttc cgt gca gca ctc gaa acc ggt gca gcc acg gta gct att tac ccc 327
Phe Arg Ala Ala Leu Glu Thr Gly Ala Ala Thr Val Ala Ile Tyr Pro
                  30                      35                      40
cgt gaa gat cgg gga tca ttc cac cgc tct ttt gct tct gaa gct gtc 375
Arg Glu Asp Arg Gly Ser Phe His Arg Ser Phe Ala Ser Glu Ala Val
                  45                      50                      55

```

cgc att ggt acc gaa ggc tca cca gtc aag gcg tac ctg gac atc gat	423
Arg Ile Gly Thr Glu Gly Ser Pro Val Lys Ala Tyr Leu Asp Ile Asp	
60 65 70 75	
gaa att atc ggt gca gct aaa aaa gtt aaa gca gat gcc att tac ccg	471
Glu Ile Ile Gly Ala Ala Lys Lys Val Lys Ala Asp Ala Ile Tyr Pro	
80 85 90	
gga tac ggc ttc ctg tct gaa aat gcc cag ctt gcc cgc gag tgt gcg	519
Gly Tyr Gly Phe Leu Ser Glu Asn Ala Gln Leu Ala Arg Glu Cys Ala	
95 100 105	
gaa aac ggc att act ttt att ggc cca acc cca gag gtt ctt gat ctc	567
Glu Asn Gly Ile Thr Phe Ile Gly Pro Thr Pro Glu Val Leu Asp Leu	
110 115 120	
acc ggt gat aag tct cgc gcg gta acc gcc gcg aag aag gct ggt ctg	615
Thr Gly Asp Lys Ser Arg Ala Val Thr Ala Ala Lys Lys Ala Gly Leu	
125 130 135	
cca gtt ttg gcg gaa tcc acc ccg agc aaa aac atc gat gag atc gtt	663
Pro Val Leu Ala Glu Ser Thr Pro Ser Lys Asn Ile Asp Glu Ile Val	
140 145 150 155	
aaa agc gct gaa ggc cag act tac ccc atc ttt gtg aag gca gtt gcc	711
Lys Ser Ala Glu Gly Gln Thr Tyr Pro Ile Phe Val Lys Ala Val Ala	
160 165 170	
ggg ggt ggc gga cgc ggt atg cgt ttt gtt gct tca cct gat gag ctt	759
Gly Gly Gly Gly Arg Gly Met Arg Phe Val Ala Ser Pro Asp Glu Leu	
175 180 185	
cgc aaa tta gca aca gaa gca tct cgt gaa gct gaa gcg gct ttc ggc	807
Arg Lys Leu Ala Thr Glu Ala Ser Arg Glu Ala Glu Ala Ala Phe Gly	
190 195 200	
gat ggc gcg gta tat gtc gaa cgt gct gtg att aac cct cag cat att	855
Asp Gly Ala Val Tyr Val Glu Arg Ala Val Ile Asn Pro Gln His Ile	
205 210 215	
gaa gtg cag atc ctt ggc gat cac act gga gaa gtt gta cac ctt tat	903
Glu Val Gln Ile Leu Gly Asp His Thr Gly Glu Val Val His Leu Tyr	
220 225 230 235	
gaa cgt gac tgc tca ctg cag cgt cgt cac caa aaa gtt gtc gaa att	951
Glu Arg Asp Cys Ser Leu Gln Arg Arg His Gln Lys Val Val Glu Ile	
240 245 250	
gcg cca gca cag cat ttg gat cca gaa ctg cgt gat cgc att tgt gcg	999
Ala Pro Ala Gln His Leu Asp Pro Glu Leu Arg Asp Arg Ile Cys Ala	
255 260 265	

gat gca gta aag ttc tgc cgc tcc att ggt tac cag ggc gcg gga acc	1047
Asp Ala Val Lys Phe Cys Arg Ser Ile Gly Tyr Gln Gly Ala Gly Thr	
270 275 280	
gtg gaa ttc ttg gtc gat gaa aag ggc aac cac gtc ttc atc gaa atg	1095
Val Glu Phe Leu Val Asp Glu Lys Gly Asn His Val Phe Ile Glu Met	
285 290 295	
aac cca cgt atc cag gtt gag cac acc gtg act gaa gaa gtc acc gag	1143
Asn Pro Arg Ile Gln Val Glu His Thr Val Thr Glu Glu Val Thr Glu	
300 305 310 315	
gtg gac ctg gtg aag gcg cag atg cgc ttg gct gct ggt gca acc ttg	1191
Val Asp Leu Val Lys Ala Gln Met Arg Leu Ala Ala Gly Ala Thr Leu	
320 325 330	
aag gaa ttg ggt ctg acc caa gat aag atc aag acc cac ggt gca gca	1239
Lys Glu Leu Gly Leu Thr Gln Asp Lys Ile Lys Thr His Gly Ala Ala	
335 340 345	
ctg cag tgc cgc atc acc acg gaa gat cca aac aac ggc ttc cgc cca	1287
Leu Gln Cys Arg Ile Thr Thr Glu Asp Pro Asn Asn Gly Phe Arg Pro	
350 355 360	
gat acc gga act atc acc gcg tac cgc tca cca ggc gga gct ggc gtt	1335
Asp Thr Gly Thr Ile Thr Ala Tyr Arg Ser Pro Gly Gly Ala Gly Val	
365 370 375	
cgt ctt gac ggt gca gct cag ctc ggt ggc gaa atc acc gca cac ttt	1383
Arg Leu Asp Gly Ala Ala Gln Leu Gly Gly Glu Ile Thr Ala His Phe	
380 385 390 395	
gac tcc atg ctg gtg aaa atg acc tgc cgt ggt tcc gac ttt gaa act	1431
Asp Ser Met Leu Val Lys Met Thr Cys Arg Gly Ser Asp Phe Glu Thr	
400 405 410	
gct gtt gct cgt gca cag cgc gcg ttg gct gag ttc acc gtg tct ggt	1479
Ala Val Ala Arg Ala Gln Arg Ala Leu Ala Glu Phe Thr Val Ser Gly	
415 420 425	
gtt gca acc aac att ggt ttc ttg cgt gcg ttg ctg cgg gaa gag gac	1527
Val Ala Thr Asn Ile Gly Phe Leu Arg Ala Leu Leu Arg Glu Glu Asp	
430 435 440	
ttc act tcc aag cgc atc gcc acc gga ttc att gcc gat cac ccg cac	1575
Phe Thr Ser Lys Arg Ile Ala Thr Gly Phe Ile Ala Asp His Pro His	
445 450 455	
ctc ctt cag gct cca cct gct gat gat gag cag gga cgc atc ctg gat	1623
Leu Leu Gln Ala Pro Pro Ala Asp Asp Glu Gln Gly Arg Ile Leu Asp	
460 465 470 475	

tac ttg gca gat gtc acc gtg aac aag cct cat ggt gtg cgt cca aag	1671
Tyr Leu Ala Asp Val Thr Val Asn Lys Pro His Gly Val Arg Pro Lys	
480 485 490	
gat gtt gca gct cct atc gat aag ctg cct aac atc aag gat ctg cca	1719
Asp Val Ala Ala Pro Ile Asp Lys Leu Pro Asn Ile Lys Asp Leu Pro	
495 500 505	
ctg cca cgc ggt tcc cgt gac cgc ctg aag cag ctt ggc cca gcc gcg	1767
Leu Pro Arg Gly Ser Arg Asp Arg Leu Lys Gln Leu Gly Pro Ala Ala	
510 515 520	
ttt gct cgt gat ctc cgt gag cag gac gca ctg gca gtt act gat acc	1815
Phe Ala Arg Asp Leu Arg Glu Gln Asp Ala Leu Ala Val Thr Asp Thr	
525 530 535	
acc ttc cgc gat gca cac cag tct ttg ctt gcg acc cga gtc cgc tca	1863
Thr Phe Arg Asp Ala His Gln Ser Leu Leu Ala Thr Arg Val Arg Ser	
540 545 550 555	
ttc gca ctg aag cct gcg gca gag gcc gtc gca aag ctg act cct gag	1911
Phe Ala Leu Lys Pro Ala Ala Glu Ala Val Ala Lys Leu Thr Pro Glu	
560 565 570	
ctt ttg tcc gtg gag gcc tgg ggc ggc gcg acc tac gat gtg gcg atg	1959
Leu Leu Ser Val Glu Ala Trp Gly Gly Ala Thr Tyr Asp Val Ala Met	
575 580 585	
cgt ttc ctc ttt gag gat ccg tgg gac agg ctc gac gag ctg cgc gag	2007
Arg Phe Leu Phe Glu Asp Pro Trp Asp Arg Leu Asp Glu Leu Arg Glu	
590 595 600	
gcg atg ccg aat gta aac att cag atg ctg ctt cgc ggc cgc aac acc	2055
Ala Met Pro Asn Val Asn Ile Gln Met Leu Leu Arg Gly Arg Asn Thr	
605 610 615	
gtg gga tac acc ccg tac cca gac tcc gtc tgc cgc gcg ttt gtt aag	2103
Val Gly Tyr Thr Pro Tyr Pro Asp Ser Val Cys Arg Ala Phe Val Lys	
620 625 630 635	
gaa gct gcc agc tcc ggc gtg gac atc ttc cgc atc ttc gac gcg ctt	2151
Glu Ala Ala Ser Ser Gly Val Asp Ile Phe Arg Ile Phe Asp Ala Leu	
640 645 650	
aac gac gtc tcc cag atg cgt cca gca atc gac gca gtc ctg gag acc	2199
Asn Asp Val Ser Gln Met Arg Pro Ala Ile Asp Ala Val Leu Glu Thr	
655 660 665	
aac acc gcg gta gcc gag gtg gct atg gct tat tct ggt gat ctc tct	2247
Asn Thr Ala Val Ala Glu Val Ala Met Ala Tyr Ser Gly Asp Leu Ser	
670 675 680	

gat cca aat gaa aag ctc tac acc ctg gat tac tac cta aag atg gca	2295
Asp Pro Asn Glu Lys Leu Tyr Thr Leu Asp Tyr Tyr Leu Lys Met Ala	
685 690 695	
gag gag atc gtc aag tct ggc gct cac atc ttg gcc att aag gat atg	2343
Glu Glu Ile Val Lys Ser Gly Ala His Ile Leu Ala Ile Lys Asp Met	
700 705 710 715	
gct ggt ctg ctt cgc cca gct gcg gta acc aag ctg gtc acc gca ctg	2391
Ala Gly Leu Leu Arg Pro Ala Ala Val Thr Lys Leu Val Thr Ala Leu	
720 725 730	
cgc cgt gaa ttc gat ctg cca gtg cac gtg cac acc cac gac act gcg	2439
Arg Arg Glu Phe Asp Leu Pro Val His Val His Thr His Asp Thr Ala	
735 740 745	
ggt ggc cag ctg gca acc tac ttt gct gca gct caa gct ggt gca gat	2487
Gly Gly Gln Leu Ala Thr Tyr Phe Ala Ala Ala Gln Ala Gly Ala Asp	
750 755 760	
gct gtt gac ggt gct tcc gca cca ctg tct ggc acc acc tcc cag cca	2535
Ala Val Asp Gly Ala Ser Ala Pro Leu Ser Gly Thr Thr Ser Gln Pro	
765 770 775	
tcc ctg tct gcc att gtt gct gca ttc gcg cac acc cgt cgc gat acc	2583
Ser Leu Ser Ala Ile Val Ala Ala Phe Ala His Thr Arg Arg Asp Thr	
780 785 790 795	
ggt ttg agc ctc gag gct gtt tct gac ctc gag ccg tac tgg gaa gca	2631
Gly Leu Ser Leu Glu Ala Val Ser Asp Leu Glu Pro Tyr Trp Glu Ala	
800 805 810	
gtg cgc gga ctg tac ctg cca ttt gag tct gga acc cca ggc cca acc	2679
Val Arg Gly Leu Tyr Leu Pro Phe Glu Ser Gly Thr Pro Gly Pro Thr	
815 820 825	
ggt cgc gtc tac cgc cac gaa atc cca ggc gga cag ttg tcc aac ctg	2727
Gly Arg Val Tyr Arg His Glu Ile Pro Gly Gly Gln Leu Ser Asn Leu	
830 835 840	
cgt gca cag gcc acc gca ctg ggc ctt gcg gat cgt ttc gaa ctc atc	2775
Arg Ala Gln Ala Thr Ala Leu Gly Leu Ala Asp Arg Phe Glu Leu Ile	
845 850 855	
gaa gac aac tac gca gcc gtt aat gag atg ctg gga cgc cca acc aag	2823
Glu Asp Asn Tyr Ala Ala Val Asn Glu Met Leu Gly Arg Pro Thr Lys	
860 865 870 875	
gtc acc cca tcc tcc aag gtt gtt ggc gac ctc gca ctc cac ctc gtt	2871
Val Thr Pro Ser Ser Lys Val Val Gly Asp Leu Ala Leu His Leu Val	
880 885 890	

ggt gcg ggt gtg gat cca gca gac ttt gct gcc gat cca caa aag tac	2919
Gly Ala Gly Val Asp Pro Ala Asp Phe Ala Ala Asp Pro Gln Lys Tyr	
895 900 905	
gac atc cca gac tct gtc atc gcg ttc ctg cgc ggc gag ctt ggt aac	2967
Asp Ile Pro Asp Ser Val Ile Ala Phe Leu Arg Gly Glu Leu Gly Asn	
910 915 920	
cct cca ggt ggc tgg cca gag cca ctg cgc acc cgc gca ctg gaa ggc	3015
Pro Pro Gly Gly Trp Pro Glu Pro Leu Arg Thr Arg Ala Leu Glu Gly	
925 930 935	
cgc tcc gaa ggc aag gca cct ctg acg gaa gtt cct gag gaa gag cag	3063
Arg Ser Glu Gly Lys Ala Pro Leu Thr Glu Val Pro Glu Glu Glu Gln	
940 945 950 955	
gcg cac ctc gac gct gat gat tcc aag gaa cgt cgc aat agc ctc aac	3111
Ala His Leu Asp Ala Asp Asp Ser Lys Glu Arg Arg Asn Ser Leu Asn	
960 965 970	
cgc ctg ctg ttc ccg aag cca acc gaa gag ttc ctc gag cac cgt cgc	3159
Arg Leu Leu Phe Pro Lys Pro Thr Glu Glu Phe Leu Glu His Arg Arg	
975 980 985	
cgc ttc ggc aac acc tct gcg ctg gat gat cgt gaa ttc ttc tac ggc	3207
Arg Phe Gly Asn Thr Ser Ala Leu Asp Asp Arg Glu Phe Phe Tyr Gly	
990 995 1000	
ctg gtc gaa ggc cgc gag act ttg atc cgc ctg cca gat gtg cgc acc	3255
Leu Val Glu Gly Arg Glu Thr Leu Ile Arg Leu Pro Asp Val Arg Thr	
1005 1010 1015	
cca ctg ctt gtt cgc ctg gat gcg atc tct gag cca gac gat aag ggt	3303
Pro Leu Leu Val Arg Leu Asp Ala Ile Ser Glu Pro Asp Asp Lys Gly	
1020 1025 1030 1035	
atg cgc aat gtt gtg gcc aac gtc aac ggc cag atc cgc cca atg cgt	3351
Met Arg Asn Val Val Ala Asn Val Asn Gly Gln Ile Arg Pro Met Arg	
1040 1045 1050	
gtg cgt gac cgc tcc gtt gag tct gtc acc gca acc gca gaa aag gca	3399
Val Arg Asp Arg Ser Val Glu Ser Val Thr Ala Thr Ala Glu Lys Ala	
1055 1060 1065	
gat tcc tcc aac aag ggc cat gtt gct gca cca ttc gct ggt gtt gtc	3447
Asp Ser Ser Asn Lys Gly His Val Ala Ala Pro Phe Ala Gly Val Val	
1070 1075 1080	
acc gtg act gtt gct gaa ggt gat gag gtc aag gct gga gat gca gtc	3495
Thr Val Thr Val Ala Glu Gly Asp Glu Val Lys Ala Gly Asp Ala Val	
1085 1090 1095	

gca atc atc gag gct atg aag atg gaa gca aca atc act gct tct gtt 3543
Ala Ile Ile Glu Ala Met Lys Met Glu Ala Thr Ile Thr Ala Ser Val
1100 1105 1110 1115

gac ggc aaa atc gat cgc gtt gtg gtt cct gct gca acg aag gtg gaa 3591
Asp Gly Lys Ile Asp Arg Val Val Val Pro Ala Ala Thr Lys Val Glu
1120 1125 1130

ggt ggc gac ttg atc gtc gtc gtt tcc taa 3621
Gly Gly Asp Leu Ile Val Val Val Ser
1135 1140

<210> 2

<211> 1140

<212> PRT

<213> Corynebacterium glutamicum

<400> 2

Met Ser Thr His Thr Ser Ser Thr Leu Pro Ala Phe Lys Lys Ile Leu
1 5 10 15

Val Ala Asn Arg Gly Glu Ile Ala Val Arg Ala Phe Arg Ala Ala Leu
20 25 30

Glu Thr Gly Ala Ala Thr Val Ala Ile Tyr Pro Arg Glu Asp Arg Gly
35 40 45

Ser Phe His Arg Ser Phe Ala Ser Glu Ala Val Arg Ile Gly Thr Glu
50 55 60

Gly Ser Pro Val Lys Ala Tyr Leu Asp Ile Asp Glu Ile Ile Gly Ala
65 70 75 80

Ala Lys Lys Val Lys Ala Asp Ala Ile Tyr Pro Gly Tyr Gly Phe Leu
85 90 95

Ser Glu Asn Ala Gln Leu Ala Arg Glu Cys Ala Glu Asn Gly Ile Thr
100 105 110

Phe Ile Gly Pro Thr Pro Glu Val Leu Asp Leu Thr Gly Asp Lys Ser
115 120 125

Arg Ala Val Thr Ala Ala Lys Lys Ala Gly Leu Pro Val Leu Ala Glu
130 135 140

Ser Thr Pro Ser Lys Asn Ile Asp Glu Ile Val Lys Ser Ala Glu Gly
145 150 155 160

Gln Thr Tyr Pro Ile Phe Val Lys Ala Val Ala Gly Gly Gly Gly Arg
165 170 175

Gly	Met	Arg	Phe	Val	Ala	Ser	Pro	Asp	Glu	Leu	Arg	Lys	Leu	Ala	Thr	180	185	190
Glu	Ala	Ser	Arg	Glu	Ala	Glu	Ala	Ala	Phe	Gly	Asp	Gly	Ala	Val	Tyr	195	200	205
Val	Glu	Arg	Ala	Val	Ile	Asn	Pro	Gln	His	Ile	Glu	Val	Gln	Ile	Leu	210	215	220
Gly	Asp	His	Thr	Gly	Glu	Val	Val	His	Leu	Tyr	Glu	Arg	Asp	Cys	Ser	225	230	235
Leu	Gln	Arg	Arg	His	Gln	Lys	Val	Val	Glu	Ile	Ala	Pro	Ala	Gln	His	245	250	255
Leu	Asp	Pro	Glu	Leu	Arg	Asp	Arg	Ile	Cys	Ala	Asp	Ala	Val	Lys	Phe	260	265	270
Cys	Arg	Ser	Ile	Gly	Tyr	Gln	Gly	Ala	Gly	Thr	Val	Glu	Phe	Leu	Val	275	280	285
Asp	Glu	Lys	Gly	Asn	His	Val	Phe	Ile	Glu	Met	Asn	Pro	Arg	Ile	Gln	290	295	300
Val	Glu	His	Thr	Val	Thr	Glu	Glu	Val	Thr	Glu	Val	Asp	Leu	Val	Lys	305	310	315
Ala	Gln	Met	Arg	Leu	Ala	Ala	Gly	Ala	Thr	Leu	Lys	Glu	Leu	Gly	Leu	325	330	335
Thr	Gln	Asp	Lys	Ile	Lys	Thr	His	Gly	Ala	Ala	Leu	Gln	Cys	Arg	Ile	340	345	350
Thr	Thr	Glu	Asp	Pro	Asn	Asn	Gly	Phe	Arg	Pro	Asp	Thr	Gly	Thr	Ile	355	360	365
Thr	Ala	Tyr	Arg	Ser	Pro	Gly	Gly	Ala	Gly	Val	Arg	Leu	Asp	Gly	Ala	370	375	380
Ala	Gln	Leu	Gly	Gly	Glu	Ile	Thr	Ala	His	Phe	Asp	Ser	Met	Leu	Val	385	390	395
Lys	Met	Thr	Cys	Arg	Gly	Ser	Asp	Phe	Glu	Thr	Ala	Val	Ala	Arg	Ala	405	410	415
Gln	Arg	Ala	Leu	Ala	Glu	Phe	Thr	Val	Ser	Gly	Val	Ala	Thr	Asn	Ile	420	425	430
Gly	Phe	Leu	Arg	Ala	Leu	Leu	Arg	Glu	Glu	Asp	Phe	Thr	Ser	Lys	Arg	435	440	445

Ile	Ala	Thr	Gly	Phe	Ile	Ala	Asp	His	Pro	His	Leu	Leu	Gln	Ala	Pro	450	455	460	
Pro	Ala	Asp	Asp	Glu	Gln	Gly	Arg	Ile	Leu	Asp	Tyr	Leu	Ala	Asp	Val	465	470	475	480
Thr	Val	Asn	Lys	Pro	His	Gly	Val	Arg	Pro	Lys	Asp	Val	Ala	Ala	Pro	485	490	495	
Ile	Asp	Lys	Leu	Pro	Asn	Ile	Lys	Asp	Leu	Pro	Leu	Pro	Arg	Gly	Ser	500	505	510	
Arg	Asp	Arg	Leu	Lys	Gln	Leu	Gly	Pro	Ala	Ala	Phe	Ala	Arg	Asp	Leu	515	520	525	
Arg	Glu	Gln	Asp	Ala	Leu	Ala	Val	Thr	Asp	Thr	Thr	Phe	Arg	Asp	Ala	530	535	540	
His	Gln	Ser	Leu	Leu	Ala	Thr	Arg	Val	Arg	Ser	Phe	Ala	Leu	Lys	Pro	545	550	555	560
Ala	Ala	Glu	Ala	Val	Ala	Lys	Leu	Thr	Pro	Glu	Leu	Leu	Ser	Val	Glu	565	570	575	
Ala	Trp	Gly	Gly	Ala	Thr	Tyr	Asp	Val	Ala	Met	Arg	Phe	Leu	Phe	Glu	580	585	590	
Asp	Pro	Trp	Asp	Arg	Leu	Asp	Glu	Leu	Arg	Glu	Ala	Met	Pro	Asn	Val	595	600	605	
Asn	Ile	Gln	Met	Leu	Leu	Arg	Gly	Arg	Asn	Thr	Val	Gly	Tyr	Thr	Pro	610	615	620	
Tyr	Pro	Asp	Ser	Val	Cys	Arg	Ala	Phe	Val	Lys	Glu	Ala	Ala	Ser	Ser	625	630	635	640
Gly	Val	Asp	Ile	Phe	Arg	Ile	Phe	Asp	Ala	Leu	Asn	Asp	Val	Ser	Gln	645	650	655	
Met	Arg	Pro	Ala	Ile	Asp	Ala	Val	Leu	Glu	Thr	Asn	Thr	Ala	Val	Ala	660	665	670	
Glu	Val	Ala	Met	Ala	Tyr	Ser	Gly	Asp	Leu	Ser	Asp	Pro	Asn	Glu	Lys	675	680	685	
Leu	Tyr	Thr	Leu	Asp	Tyr	Tyr	Leu	Lys	Met	Ala	Glu	Glu	Ile	Val	Lys	690	695	700	
Ser	Gly	Ala	His	Ile	Leu	Ala	Ile	Lys	Asp	Met	Ala	Gly	Leu	Leu	Arg	705	710	715	720
Pro	Ala	Ala	Val	Thr	Lys	Leu	Val	Thr	Ala	Leu	Arg	Arg	Glu	Phe	Asp				

	725		730		735										
Leu	Pro	Val	His	Val	His	Thr	His	Asp	Thr	Ala	Gly	Gly	Gln	Leu	Ala
		740						745					750		
Thr	Tyr	Phe	Ala	Ala	Ala	Gln	Ala	Gly	Ala	Asp	Ala	Val	Asp	Gly	Ala
		755						760					765		
Ser	Ala	Pro	Leu	Ser	Gly	Thr	Thr	Ser	Gln	Pro	Ser	Leu	Ser	Ala	Ile
		770						775				780			
Val	Ala	Ala	Phe	Ala	His	Thr	Arg	Arg	Asp	Thr	Gly	Leu	Ser	Leu	Glu
		785								795					800
Ala	Val	Ser	Asp	Leu	Glu	Pro	Tyr	Trp	Glu	Ala	Val	Arg	Gly	Leu	Tyr
				805					810						815
Leu	Pro	Phe	Glu	Ser	Gly	Thr	Pro	Gly	Pro	Thr	Gly	Arg	Val	Tyr	Arg
			820						825					830	
His	Glu	Ile	Pro	Gly	Gly	Gln	Leu	Ser	Asn	Leu	Arg	Ala	Gln	Ala	Thr
		835						840					845		
Ala	Leu	Gly	Leu	Ala	Asp	Arg	Phe	Glu	Leu	Ile	Glu	Asp	Asn	Tyr	Ala
		850				855					860				
Ala	Val	Asn	Glu	Met	Leu	Gly	Arg	Pro	Thr	Lys	Val	Thr	Pro	Ser	Ser
		865			870					875					880
Lys	Val	Val	Gly	Asp	Leu	Ala	Leu	His	Leu	Val	Gly	Ala	Gly	Val	Asp
			885						890					895	
Pro	Ala	Asp	Phe	Ala	Ala	Asp	Pro	Gln	Lys	Tyr	Asp	Ile	Pro	Asp	Ser
		900						905					910		
Val	Ile	Ala	Phe	Leu	Arg	Gly	Glu	Leu	Gly	Asn	Pro	Pro	Gly	Gly	Trp
		915					920					925			
Pro	Glu	Pro	Leu	Arg	Thr	Arg	Ala	Leu	Glu	Gly	Arg	Ser	Glu	Gly	Lys
		930				935					940				
Ala	Pro	Leu	Thr	Glu	Val	Pro	Glu	Glu	Glu	Gln	Ala	His	Leu	Asp	Ala
		945			950					955					960
Asp	Asp	Ser	Lys	Glu	Arg	Arg	Asn	Ser	Leu	Asn	Arg	Leu	Leu	Phe	Pro
			965						970					975	
Lys	Pro	Thr	Glu	Glu	Phe	Leu	Glu	His	Arg	Arg	Arg	Phe	Gly	Asn	Thr
		980						985					990		
Ser	Ala	Leu	Asp	Asp	Arg	Glu	Phe	Phe	Tyr	Gly	Leu	Val	Glu	Gly	Arg
		995				1000						1005			

Glu Thr Leu Ile Arg Leu Pro Asp Val Arg Thr Pro Leu Leu Val Arg
1010 1015 1020

Leu Asp Ala Ile Ser Glu Pro Asp Asp Lys Gly Met Arg Asn Val Val
1025 1030 1035 1040

Ala Asn Val Asn Gly Gln Ile Arg Pro Met Arg Val Arg Asp Arg Ser
1045 1050 1055

Val Glu Ser Val Thr Ala Thr Ala Glu Lys Ala Asp Ser Ser Asn Lys
1060 1065 1070

Gly His Val Ala Ala Pro Phe Ala Gly Val Val Thr Val Thr Val Ala
1075 1080 1085

Glu Gly Asp Glu Val Lys Ala Gly Asp Ala Val Ala Ile Ile Glu Ala
1090 1095 1100

Met Lys Met Glu Ala Thr Ile Thr Ala Ser Val Asp Gly Lys Ile Asp
1105 1110 1115 1120

Arg Val Val Val Pro Ala Ala Thr Lys Val Glu Gly Gly Asp Leu Ile
1125 1130 1135

Val Val Val Ser
1140

<210> 3
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward DNA
Primer

<400> 3
gtcttcacgc agatgaatcc gcg

23

<210> 4
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Reverse DNA
Primer

<400> 4

cgcagcgcca catcgtaagt cgc

23

<210> 5

<211> 8

<212> PRT

<213> *Caenorhabditis elegans*

<400> 5

Tyr Phe Ile Glu Val Asn Ala Arg

1

5

<210> 6

<211> 7

<212> PRT

<213> *Caenorhabditis elegans*

<400> 6

Ala Thr Phe Asp Val Ser Met

1

5

<210> 7

<211> 8

<212> PRT

<213> *Aedes aegypti*

<400> 7

Tyr Phe Ile Glu Val Asn Ala Arg

1

5

<210> 8

<211> 7

<212> PRT

<213> *Aedes aegypti*

<400> 8

Ala Thr Phe Asp Val Ala Leu

1

5

<210> 9

<211> 8

<212> PRT

<213> *Mycobacterium tuberculosis*

<400> 9

Val Phe Ile Glu Met Asn Pro Arg

1

5

<210> 10
<211> 7
<212> PRT
<213> Mycobacterium tuberculosis

<400> 10
Ala Thr Tyr Asp Val Ala Leu
1 5

<210> 11
<211> 8
<212> PRT
<213> Bacillus stearothermophilus

<400> 11
Tyr Phe Ile Glu Val Asn Pro Arg
1 5

<210> 12
<211> 7
<212> PRT
<213> Bacillus stearothermophilus

<400> 12
Ala Thr Phe Asp Val Ala Tyr
1 5

<210> 13
<211> 8
<212> PRT
<213> Pichia pastoris

<400> 13
Tyr Phe Ile Glu Ile Asn Pro Arg
1 5

<210> 14
<211> 7
<212> PRT
<213> Pichia pastoris

<400> 14
Ala Thr Phe Asp Val Ser Met
1 5

<210> 15

<211> 8
<212> PRT
<213> Mus musculus

<400> 15
Tyr Phe Ile Glu Val Asn Ser Arg
1 5

<210> 16
<211> 7
<212> PRT
<213> Mus musculus

<400> 16
Ala Thr Phe Asp Val Ala Met
1 5

<210> 17
<211> 8
<212> PRT
<213> Rattus norvegicus

<400> 17
Tyr Phe Ile Glu Val Asn Ser Arg
1 5

<210> 18
<211> 7
<212> PRT
<213> Rattus norvegicus

<400> 18
Ala Thr Phe Asp Val Ala Met
1 5

<210> 19
<211> 8
<212> PRT
<213> Saccharomyces cerevisiae 1

<400> 19
Tyr Phe Ile Glu Ile Asn Pro Arg
1 5

<210> 20
<211> 7
<212> PRT

<213> Saccharomyces cerevisiae 1

<400> 20

Ala Thr Phe Asp Val Ala Met

1

5

<210> 21

<211> 8

<212> PRT

<213> Saccharomyces cerevisiae 2

<400> 21

Tyr Phe Ile Glu Ile Asn Pro Arg

1

5

<210> 22

<211> 7

<212> PRT

<213> Saccharomyces cerevisiae 2

<400> 22

Ala Thr Phe Asp Val Ala Met

1

5

<210> 23

<211> 8

<212> PRT

<213> Rhizobium etli

<400> 23

Tyr Phe Ile Glu Val Asn Pro Arg

1

5

<210> 24

<211> 7

<212> PRT

<213> Rhizobium etli

<400> 24

Ala Thr Phe Asp Val Ser Met

1

5

<210> 25

<211> 8

<212> PRT

<213> Homo sapiens

<400> 25

Tyr Phe Ile Glu Val Asn Ser Arg
1 5

<210> 26

<211> 7

<212> PRT

<213> Homo sapiens

<400> 26

Ala Thr Phe Asp Val Ala Met
1 5

<210> 27

<211> 8

<212> PRT

<213> Schizosaccharomyces pombe

<400> 27

Tyr Phe Ile Glu Ile Asn Pro Arg
1 5

<210> 28

<211> 7

<212> PRT

<213> Schizosaccharomyces pombe

<400> 28

Ala Thr Phe Asp Val Ser Met
1 5

<210> 29

<211> 9

<212> PRT

<213> Corynebacterium glutamicum

<400> 29

Phe Leu Phe Glu Asp Pro Trp Asp Arg
1 5

<210> 30

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: DNA Primer

bioRxiv preprint doi: <https://doi.org/10.1101/201904.01.201904>; this version posted April 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

<400> 30
ttcaccaggt ccacctcg 18

<210> 31
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 31
cgtcgcaaag ctgactcc 18

<210> 32
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 32
gatgcttctg ttgctaattt gc 22

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 33
ggccattaag gatatggctg 20

<210> 34
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 34
gcggtggaat gatccccga 19

<210> 35
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 35
accgcactgg gccttgcg 18

<210> 36
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA Primer

<400> 36
tcgccgcttc ggcaacac 18

<210> 37
<211> 6
<212> PRT
<213> Corynebacterium glutamicum

<400> 37
Gly Gly Gly Gly Arg Gly
1 5

<210> 38
<211> 6
<212> PRT
<213> Corynebacterium glutamicum

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> May be Arg or Lys

<400> 38
Gly Gly Gly Gly Xaa Gly
1 5

<210> 39
<211> 4
<212> PRT
<213> Unknown

<220>

<223> Putative biotin-binding site

<400> 39

Ala Met Lys Met

1

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.